

Modeling for evolving biological networks with scale-free connectivity, hierarchical modularity, and disassortativity

Kazuhiro Takemoto^{a,*}, Chikoo Oosawa^{a,b,**}

^a*Department of Bioscience and Bioinformatics, Kyushu Institute of Technology,
Izuka, Fukuoka 820-8502, Japan*

^b*Bioalgorithm Project, Faculty of Computer Science and Systems Engineering,
Kyushu Institute of Technology, Izuka, Fukuoka 820-8502, Japan*

Abstract

We propose a growing network model that consists of two tunable mechanisms: growth by merging modules which are represented as complete graphs and a fitness-driven preferential attachment. Our model exhibits the three prominent statistical properties are widely shared in real biological networks, for example gene regulatory, protein-protein interaction, and metabolic networks. They retain three power law relationships, such as the power laws of degree distribution, clustering spectrum, and degree-degree correlation corresponding to scale-free connectivity, hierarchical modularity, and disassortativity, respectively. After making comparisons of these properties between model networks and biological networks, we confirmed that our model has inference potential for evolutionary processes of biological networks.

Key words: power law; biological networks; scale free; hierarchical modularity; disassortativity

1 Introduction

Structural data of large-scale biological networks such as gene regulatory networks [1,2,3], protein-protein interaction networks [4,5,6,7], and metabolic networks [8,9] have been accumulated by the active investigations in recent years,

* Current affiliation: Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan.

**Corresponding author.

Email address: `chikoo@bio.kyutech.ac.jp` (Chikoo Oosawa).

and heterogeneous connectivity, which means that a few nodes integrate a great number of nodes and most of the remaining nodes do not, are ubiquitously found as a striking property of the structures [2,6,7,8,9,10,11,12,13,14]. Heterogeneity exhibits the scale-free feature [15,16], and differs from homogeneous connectivity, which means that most nodes have the same degree (number of edges), is observed in classical network models such as regular lattices and the Erdős-Rényi random network [17]. The scale-free feature is in degree distribution, defined as the probability of the existence nodes with degree k , which follows the power law, $P(k) \propto k^{-\gamma}$ with $2 < \gamma < 3$. Since the scale-free feature is independent of living species [8,11,12,18], growth mechanism which can generate such statistical properties are paid much attention, and are expected to elucidate the evolutionary and growing processes of the networks.

In previous works, the Barabási-Albert (BA) model [16] is well-known as a model for scale-free networks, and is frequently embedded among variety of mathematical models [19]. Using the growth mechanism by adding new nodes with preferential attachment (PA) which means that high-degree nodes get an even better chance to attract new edges, the BA model generates scale-free networks which indicate power-law degree distributions with the fixed degree exponent -3 ; thus $P(k) \propto k^{-3}$ [16,18,20]. However, the recent detailed statistical analyses of biological networks manifest some discrepancies between the networks generated by the BA model and biological networks. In particular, there are two striking properties as follows.

One of the properties is a small-world feature [21]. This feature is reflected in high clustering coefficients C , which denote the density of edges between neighbors of a given node, and implies modularity of networks [22,23]. The modular structures are actively investigated with statistical approaches, and it is found that clustering spectra, defined as correlations between degree k of a given node and the clustering coefficient C of the node, follow the power law with exponent -1 ; thus $C(k) \propto k^{-1}$ [13,14,24]. The power-law spectrum, observed in biological networks, suggests a hierarchical structure of modules [24]. However, BA networks do not exhibit a power-law clustering spectrum, indicating the absence of the hierarchical structures [25,26]. In order to fill the gap, we propose a model (hereinafter called MM model) with growth mechanism by merging modules and PA from the BA model. The MM model shows power-law clustering spectra and power-law degree distributions with arbitrary degree exponents, demonstrating that our model can reproduce finer details of biological networks [27].

The other of the properties is a heterogeneous degree-degree correlation [28], defined as correlation between degree k of a given node and average degree \bar{k}_{nn} of neighbors of the node. In biological networks, the degree-degree correlations follow power law; $\bar{k}_{nn}(k) \propto k^\nu$ with a negative exponent, $-1 < \nu < 0$, empirically found [13,14,29]. In general, the properties are called disassortativity

[30]. It is reported that the BA and MM model have no disassortativity [27,28]. Pastor-Satorras *et al.* [28] and Barrat *et al.* [31] respectively argue that disassortativity emerges with competitive dynamics [32] and weight-driven dynamics [33], suggesting that growing mechanism needs to include fitness such as spatial distance [34], aging [35], and so on.

In this paper, we propose a model which maintains the following statistical properties: the scale-free feature, the hierarchical modularity, and the disassortativity. For the proposal, we start with a modification of the MM model that has the scale-free feature and the hierarchical modularity [27]. We embed a tunable fitness-driven (FD) mechanism into the MM model, and provide the numerical and analytical solutions for the statistical properties of model networks.

2 Model

Here we present an evolving network model. Our model includes growth by merging modules and the FD mechanism [see Section 10 for biological implications of the model].

Growth by merging modules — A network grows through joining new modules to existing nodes of a network over time [see Fig. 1 (C)]. m/a is the first control parameter of the model where a and m denote the number of nodes of the complete graph and the number of merged node(s), respectively. Please note that the process develops without adding extra edges [27].

Fitness-driven (FD) preferential attachment (PA) — The standard PA mechanism of the BA model is the probability that node i is chosen to get an edge, and is proportional to the degree of node i ; hence $\Pi_i^{\text{BA}} = k_i / \sum_j k_j$, where k_i is the degree of node i . The mechanism only considers the degrees at the nodes. Here, we additionally consider the probability that node i is selected according to degree k_i and fitness f_i , and express the probability as

$$\Pi_i = \frac{k_i + f_i}{\sum_j (k_j + f_j)}. \quad (1)$$

Updating rule of fitness — Moreover, we consider the change of the fitness of node i . When node i is selected by using the FD-PA mechanism given in Eq. (1), the fitness of node i increases as follows:

$$f_i \leftarrow f_i + \xi, \quad (2)$$

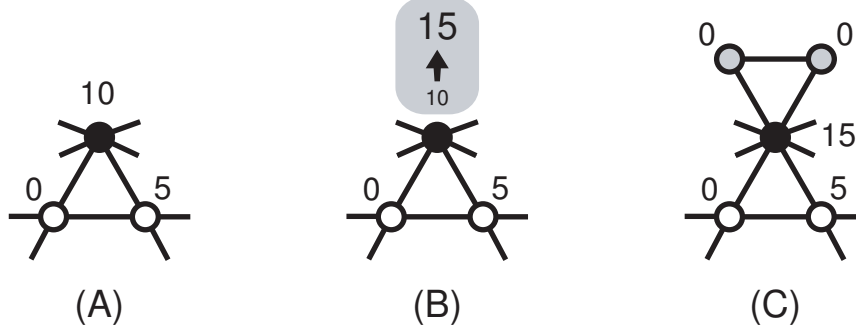


Fig. 1. Schematic diagram of growth process of our model network with $a = 3$, $m = 1$, and $\xi = 5$. (A) Selection of node(s) by the FD-PA mechanism, Eq. (1). The filled node is selected by the attachment. Each number in the figure indicates the fitness of a corresponding node. (B) Updating of the fitness. The selected node's fitness increases according to the updating rule, Eq. (2). (C) Merging new modules. As a result, $(a - m)$ new node(s) are added, filled with gray, and take zero as their initial fitness.

where ξ takes a constant value, and is the second control parameter in the model, indicating the strength of incidence of the fitness.

Taken together, our model networks are generated by the following procedures.

- i) We start from a module that is a complete graph consisting of a (≥ 3) nodes. For fitness, we assign zero to all nodes in the module.
- ii) At every time step a new module with the same size is joined by merging to an existing m ($< a$) nodes.
- iii) When merging the module, the FD-PA mechanism, Eq (1), is used to select m old nodes [see Fig. 1 (A)]. Then, the fitnesses of the old nodes are increased using the updating rule, Eq. (2) [see Fig. 1 (B)]. At the end, zero is assigned to the fitnesses of the new nodes [see Fig. 1 (C)]. Resultant duplicated edges between the merged nodes are counted and contribute to the FD-PA mechanism in the succeeding steps.

When $a > m$ and $\xi \geq 0$, the model network evolves in time steps. In addition, our model is equivalent to the BA model with the specific condition: $a = 2$, $m = 1$, and $\xi = 0$.

3 Degree distribution

In this section, we present the analytical and the numerical solutions for the degree distribution of our model. The degree distribution is an important statistical property for the characterization of the networks, and denotes the existence probability of nodes with degree k . The degree distribution [18,20]

is defined as

$$P(k) = \frac{1}{N} \sum_{i=1}^N \delta(k_i - k), \quad (3)$$

where $\delta(x)$ and N are Kronecker's delta function and the total number of the nodes, respectively.

3.1 Analytical solution

In order to describe the degree distribution, we take the continuous mean-field approach [16,26]. The standard approach can not be applied directly due to the inclusion of the fitness updating. We express the degree and the fitness as

$$F_i = k_i + f_i, \quad (4)$$

and investigate a time evolution of F_i . Since the fitness updating rule is represented as $f_i \leftarrow f_i + \xi$, F_i satisfies

$$F_i = \left(\frac{\xi}{a-1} + 1 \right) k_i - \xi, \quad (5)$$

indicating the proportional relationship between F_i and k_i .

The time evolution of F_i is described as

$$\frac{dF_i}{dt} = m(a-1+\xi) \frac{F_i}{\sum_j F_j}, \quad (6)$$

where $\sum_j F_j \approx [a(a-1) + m\xi]t$. The solution of the equation with $F_i(t=s) = \binom{a}{2} + \xi = A(a, \xi)$ as an initial condition for Eq. (6) is

$$F_i(t) = A(a, \xi) \left(\frac{t}{s} \right)^\beta, \quad (7)$$

where $\beta = [m(a-1+\xi)]/[a(a-1) + m\xi]$. Because s/t denotes the probability that F_i is larger than a given F , Equation (7) is rewritten as

$$P(> F) = A(a, \xi)^{1/\beta} F^{-1/\beta}, \quad (8)$$

which corresponds to the cumulative probability. From Eq. (8), the probability distribution for F is given as

$$P(F) = -\frac{d}{dF}P(> F) = \frac{A(a, \xi)^{1/\beta}}{\beta} F^{-\gamma}, \quad (9)$$

where $\gamma = (1/\beta) + 1$. Finally, substituting Eq. (5) into Eq. (9), we get the degree distribution

$$P(k) \simeq B(a, \xi, \gamma) k^{-\gamma}, \quad (10)$$

where $B(a, \xi, \gamma) = A(a, \xi)^{\gamma-1} [\xi/(a-1) + 1]^{-\gamma}/(\gamma-1)$. The degree distribution obeys the power-law with the degree exponent

$$\gamma = \frac{(a+m)(a-1) + 2m\xi}{m(a-1+\xi)}, \quad (11)$$

demonstrating that our model network represents the scale-free feature.

3.2 Numerical solution

In order to confirm the analytical predictions, we performed numerical simulations of networks generated by using our model described in Sec. 2. We show degree distributions with $a = 5$, $N = 50000$, different $m \in [1, 4]$ and $\xi \in [0, 15]$ in Fig. 2. The degree distributions follow the power law, reflecting the scale-free feature.

Figure 2 (A) shows the degree distributions with $m = 2$ and the different ξ . The degree exponents decay with increasing ξ . Figure 2 (B) shows the degree distributions with $\xi = 7$ and the different m . The degree exponents decay with increasing m . We show excellent agreement between the numerical results and the theoretical predictions, demonstrating that γ is function of m/a and ξ .

4 Degree entropy

As shown in Eq. (11) and Fig. 2, the degree exponent γ is variant with ξ and m/a . We consider the degree entropy [36] to characterize parameter dependency for the heterogeneous connectivity of our model. Assuming that

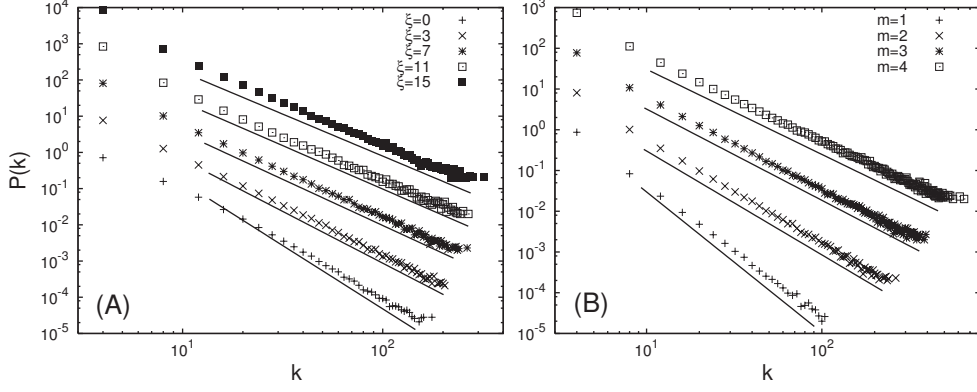


Fig. 2. Degree distributions $P(k)$ with $a = 5$ and $N = 50000$ (shifted for clarity). Different symbols correspond to the different numerical results. Solid lines show relationship $\propto k^{-\gamma}$, where γ is predicted by Eq. (11). (A) ξ dependency with $m = 2$. (B) m/a dependency with $\xi = 7$.

$\sum_{k=1}^{N-1} P(k) = 1$, the degree entropy is defined as

$$H = - \sum_{k=1}^{N-1} P(k) \ln P(k), \quad (12)$$

and provides an average measure of a network's heterogeneity because it relies on the diversity of the degree distribution. The increment of degree entropy H corresponds to the increment of the heterogeneity for complex networks [36].

Figure 3 shows the numerical results of the degree entropy of the model network with $a = 5$, $N = 12800$, different $m \in [1, 4]$, and different $\xi \in [0, 15]$.

In Fig 3 (A), we show the degree entropy H with the different m/a and ξ in three dimensions. The degree entropy H increases for the larger m/a and the smaller ξ , and decreases in the opposite direction. In Fig 3 (B), the projections of the plot to the (m/a) - H plane of Fig. 3 (A) are shown. The correlation between H and m/a becomes strong with increasing ξ . In Fig 3 (C), the projections of the plot to the ξ - H plane of Fig. 3 (A) are shown. The degree entropy decays with increasing ξ in the case of constant m/a .

5 Clustering spectrum

The clustering spectrum is a well-known statistical property which reflects the hierarchical modularity of networks [18,20], and is defined as

$$C(k) = \frac{\sum_{i=1}^N C_i \times \delta(k - k_i)}{NP(k)}, \quad (13)$$

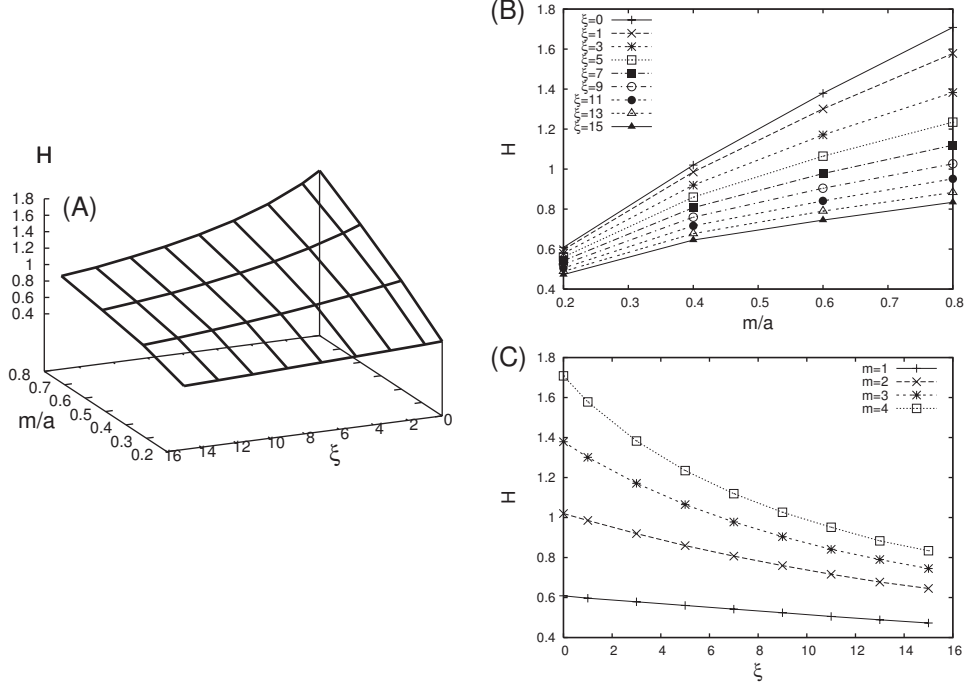


Fig. 3. Degree entropy H with $a = 5$ and $N = 12800$ (Axis direction changed for clarity). (A) 3D plot of the entropy versus the variables ξ and m/a . Projections of the plot to the (m/a) - H plane (B) and the ξ - H plane (C).

where $\delta(x)$ denotes Kronecker's delta function, and C_i corresponds to the clustering coefficient [21], defined as

$$C_i = \frac{M_i}{\binom{k_i}{2}} = \frac{2M_i}{k_i(k_i - 1)}, \quad (14)$$

and represents the density of edges among neighbors of a given node, where k_i and M_i denote the degree of node i and the number of edges among the neighbors, respectively.

First, we give the analytical solution for the clustering spectrum of our model. Since our model grows through merging modules (see Sec. 2 and Fig. 1), the number of edges of node i among neighboring nodes is approximately described as [27]

$$M_i \simeq S_i \binom{a-1}{2} = S_i(a-1) \frac{a-2}{2}, \quad (15)$$

where S_i corresponds to the number of the selection of node i with the PA, as in Eq (1). In our model, because the degree of node i is expressed as

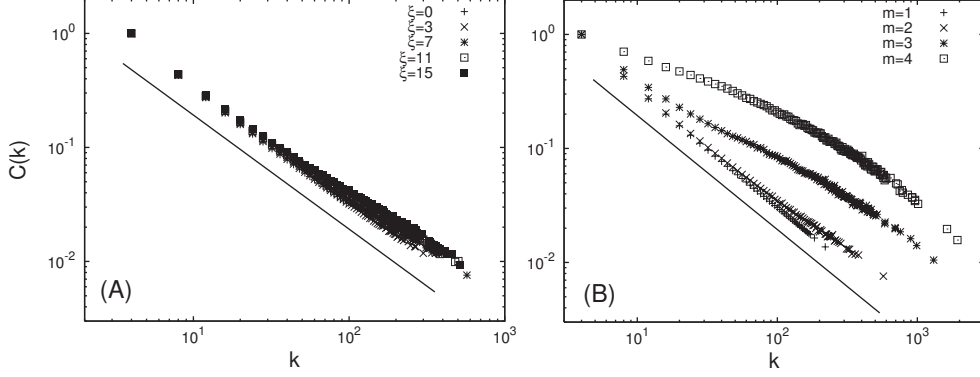


Fig. 4. Cluster spectra $C(k)$ with $a = 5$ and $N = 50000$. Different symbols correspond to the numerical results. Solid lines show the relationship $\propto k^{-1}$. (A) ξ dependency with $m = 2$. (B) m dependency with $\xi = 7$.

$k_i = S_i(a - 1)$, Eq. (15) is rewritten as

$$M_i \simeq \frac{a-2}{2} k_i, \quad (16)$$

indicating the proportional relationship between M_i and k_i . Finally, substituting Eq. (16) into Eq. (14), we get the clustering spectrum

$$C(k) \simeq \frac{a-2}{k} \propto k^{-1}. \quad (17)$$

The clustering spectrum follows the power law with the exponent -1 when $a \geq 3$, reflecting the hierarchical modularity of our model network.

Next, we present the numerical results of our model for the clustering spectrum to verify the analytical solution. In Fig. 4, we show clustering spectra with $a = 5$, $N = 50000$, different $m \in [1, 4]$, and different $\xi \in [0, 15]$.

Figure 4 (A) shows the clustering spectra with $m = 2$ and the different ξ . The spectra follow power law with the exponent -1 despite changing ξ . Figure 4 (B) shows the clustering spectra with $\xi = 7$ and the different m . The spectra follows power law with the exponent -1 for large k . For small degree k , the cut-off is prominent with increasing m due to running away from the approximation of Eq. (15).

The clustering spectra follow the power law with the exponent more or less equal to -1 , reflecting the hierarchical modularity of networks.

6 Global clustering coefficient

As seen in Fig. 3, degree entropy H is sensitive to ξ and m/a . To examine parameter-dependency of the modularity (clustering property) for the whole network, we utilize a global clustering coefficient C . The coefficient is defined as [21]

$$C = \frac{1}{N} \sum_{i=1}^N C_i. \quad (18)$$

A higher value of C means that the network has higher modularity.

Figure 5 shows the numerical result of C in our model network with $a = 5$, $N = 12800$, different $m \in [1, 4]$, and different $\xi \in [0, 15]$.

In Fig. 5 (A), we show the C with different m/a and ξ in three dimensions. C tends to increase more for smaller m/a and larger ξ .

Since the small or large m/a give rise to different effects on C , Figure 5 (A) has a valley in the middle of m/a . In the case of small m/a , C can remain high because the modules combine via few common nodes. In the case of large m/a , in contrast, C leniently decreases because the networks tend to be randomized since the modules combine via many common nodes. In this case, however, FD mechanism helps to increase the fitness of the nodes in the modules, inducing a formation of a cluster with high-edge density. As a result, C increases with m/a . Due to the trade-off mechanism, the biphasic graph is shown in Fig. 5 (B). In addition, C is less sensitive for large ξ . In Fig. 5 (C), we show the projection of the plot to the ξ - C plane of Fig. 5 (A).

7 Degree-degree correlation

The degree-degree correlation is a statistical property that characterizes assortativity of the networks, and represents the average degree of neighbors of nodes with degree k [18,20]. The correlation is defined as

$$\bar{k}_{nn}(k) = \sum_{k'} k' P(k'|k), \quad (19)$$

where the conditional probability $P(k'|k)$ is the frequency that a node with degree k connects to a node with degree k' . Using Kronecker's delta function,

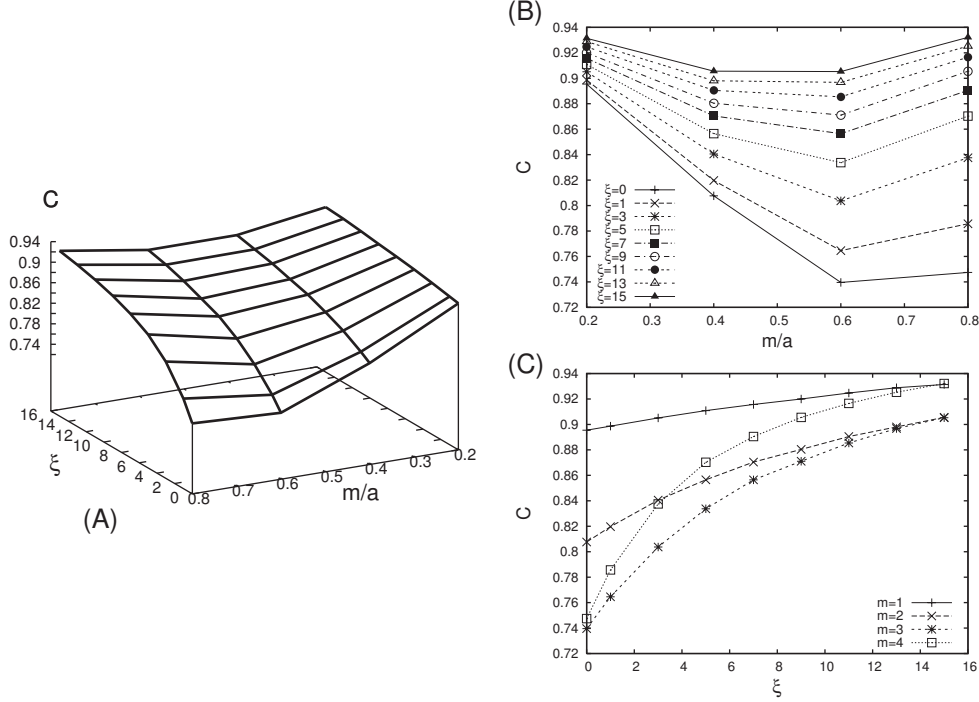


Fig. 5. Global clustering coefficient C with $a = 5$ and $N = 12800$. (A) 3D plot of the coefficient versus the variables, ξ and m/a (axis direction changed for clarity). Projections of the plot to the (m/a) - C plane (B) and the ξ - C plane (C).

we redefined the degree-degree correlation as

$$\bar{k}_{nn}(k) = \frac{\sum_{i=1}^N \langle k_{nn} \rangle_i \times \delta(k_i - k)}{NP(k)}, \quad (20)$$

where $\langle k_{nn} \rangle_i$ denotes the average nearest-neighbor degree, written as

$$\langle k_{nn} \rangle_i = \frac{1}{k_i} \sum_{h \in V(i)} k_h, \quad (21)$$

where $V(i)$ corresponds to the set of neighbors of node i .

Here, we give the numerical solutions for the degree-degree correlation of our model. Figure 6 shows the degree-degree correlations with $a = 5$, $N = 50000$, the different m , and ξ . The correlations follow the power law; this $\bar{k}_{nn}(k) \propto k^\nu$ with $-1 < \nu < 0$ as a roughly observation. Negative and larger ν tend to be seen for the larger m/a and ξ .

Figure 6 (A) shows the degree-degree correlations with $m = 2$ and the different ξ . The exponent ν decays with increasing ξ . Figure 6 (A) shows the degree-degree correlations with $\xi = 7$ and the different m . The exponent ν decays with increasing m .

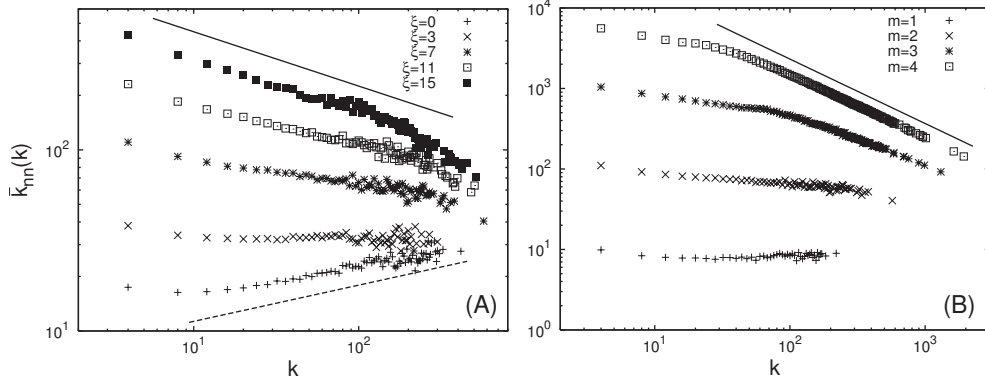


Fig. 6. Degree correlations $\bar{k}_{nn}(k)$ with $a = 5$ and $N = 50000$. (A) ξ dependency with $m = 2$. A solid and dashed line correspond to $\propto k^{-0.3}$ and $\propto k^{0.2}$, respectively. (B) m dependency with $\xi = 7$. A solid line shows the relationship $\propto k^{-0.8}$.

Due to the fitness updating and the PA, the degree-degree correlations follow the power law, reflecting the disassortativity of the networks. As previously reported, the disassortativity is not reproduced if we consider the PA mechanism only [27,28].

8 Assortative coefficient

The assortative coefficient (AC) [30] can be thought of as a compendium parameter of the degree-degree correlations, and is defined as

$$r = \frac{4\langle k_i k_j \rangle - \langle k_i + k_j \rangle^2}{2\langle k_i^2 + k_j^2 \rangle - \langle k_i + k_j \rangle^2} \quad (22)$$

where k_i and k_j are the degrees of two vertices at the ends of an edge, and $\langle \dots \rangle$ denotes the average over all edges. In other words, the AC is the correlation coefficient for the degree-degree correlation $\bar{k}_{nn}(k)$, and takes values $-1 \leq r \leq 1$. The relationship between the AC and network structures is described as follows:

- i) In the case of $r < 0$, low-degree nodes tend to connect to high-degree nodes, indicating the disassortativity. Then, the degree-degree correlation $\bar{k}_{nn}(k)$ decreases with increasing k .
- ii) In the case of $r = 0$, the degree-degree correlation $\bar{k}_{nn}(k)$ is uncorrelated.
- iii) In the case of $r > 0$, high-degree nodes tend to connect to high-degree nodes, reflecting the assortativity. Then, the degree-degree correlation $\bar{k}_{nn}(k)$ increases with degree k .

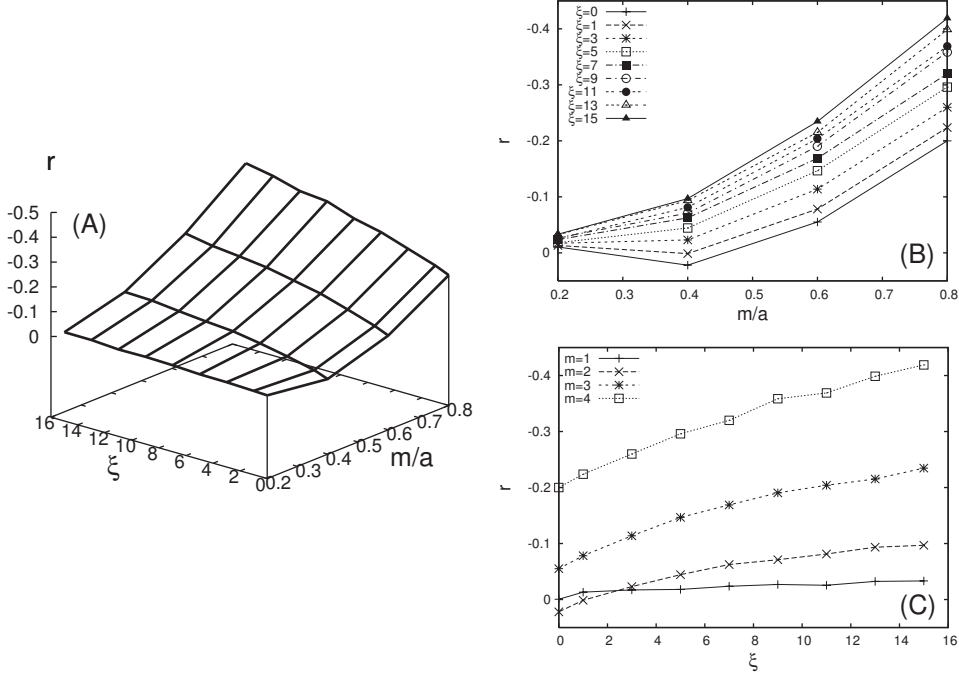


Fig. 7. Assortative coefficient r with $a = 5$ and $N = 12800$. (A) 3D plot of the coefficient versus the variables, ξ and m/a . Projections of the plot to the (m/a) - r plane (B) and the ξ - r plane (C). Vertical axis of (A)–(C) is inverted for clarity.

Figure 7 shows the numerical solutions of the AC for our model with $a = 5$, $N = 12800$, different $m \in [1, 4]$, and $\xi \in [0, 15]$.

Figure 7 (A) shows the AC with the different m/a and ξ in three dimensions. The large negative AC generally tends to be seen for the larger m/a and ξ . Figure 7 (B) shows the projections of the plot to the (m/a) - r plane of Fig. 7 (A). Two biphasic curves (once r takes positive instead of negative) for the smaller $\xi \in [0, 1]$ emerge because positive degree-degree correlations are exhibited in the case of $m/a \leq 0.5$ [27]. For larger $\xi \in [3, 15]$, assortative coefficients r monotonically decrease (upward) with increasing m/a . Figure 7 (C) shows the projections of the plot on the ξ - r plane for Fig. 7 (A).

9 Comparison of statistical properties

In order to validate our model we make a comparison of statistical properties between biological networks and our model networks. We prepare two different types of networks; the gene regulatory [37] and the metabolic [8] networks of *Escherichia coli*. The gene regulatory network is represented as sets of graphs consisting of nodes and edges; they correspond to genes and interactions among genes, respectively. For simplicity, we extract the largest component from the networks, and replace directed and/or weighted edges

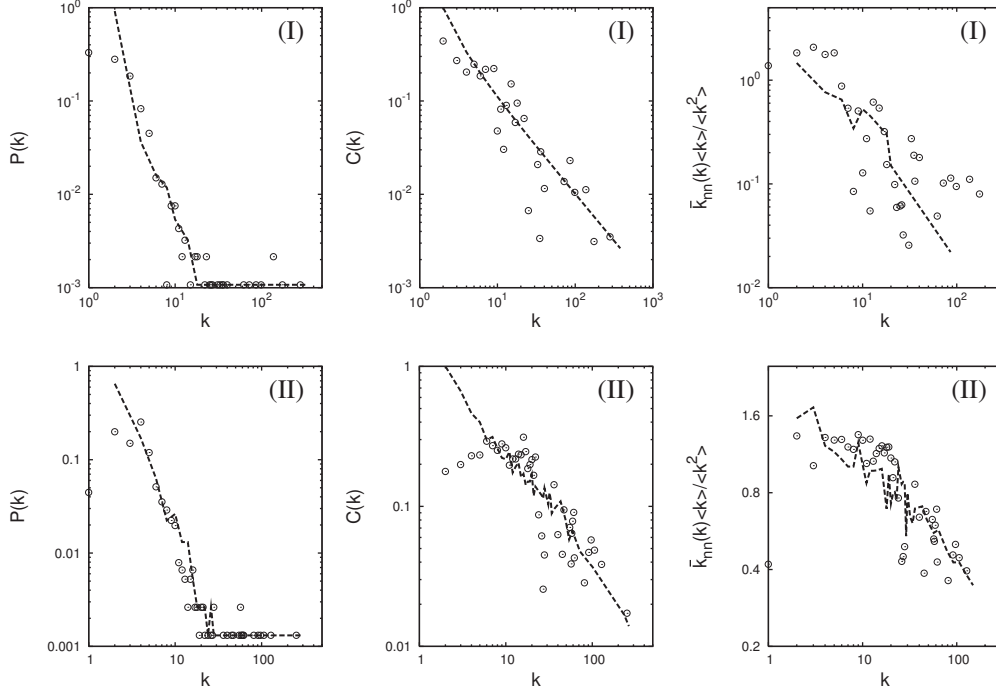


Fig. 8. Comparison of statistical properties between biological networks and model networks. The left column shows degree distributions $P(k)$. Clustering spectra $C(k)$ are in center column. Degree-degree correlations $\bar{k}_{nn}(k)$ constitute the right column. Note that the degree-degree correlations $\bar{k}_{nn}(k)$ are divided by $\langle k^2 \rangle / \langle k \rangle$ for normalization. Symbols and dashed lines denote data of biological networks and our model networks, respectively. (I) Gene regulatory network in *Escherichia coli* [37] and our model with $a = 3$, $m = 1$, and $\xi = 20$. (II) Metabolic networks in *Escherichia coli* [8] and our model with $a = 3$, $m = 2$, and $\xi = 1$. The size of the model network is the same as total number of nodes of the biological network.

with undirected and unweighted edges. Moreover, we remove multiple edges and self-loops. The metabolic network is transformed with the same procedures.

In Fig. 8, we show the statistical properties of the biological networks and our model networks. We have good agreement with the real network's data and our model, demonstrating that our model reproduces the three striking statistical properties, also widely shared in the biological networks. In addition we confirm similar relationships with network data derived from the yeast [38] and the worm [39] (data not shown).

10 Discussions

Even though large amounts of structural data of biological networks have been available in recent years, little is known about the detailed mechanisms

behind the origin and evolution of the networks. However, we proposed a hypothetical model and its parameter space that can reproduce biological features from a statistical viewpoint. We think that the growth mechanism by merging modules and the FD-PA mechanism are important for evolutionary processes in biological networks.

Since new genes commonly appear through complete or partial duplication of pre-existing genes [40,41,42,43,44], the resultant sequences are subject to selective pressure as parts of the whole system and would give rise to changes in functional interactions among biomolecules, resulting in changes of statistical properties of network topology.

In the case of metabolic networks, modules in our model can be regarded as a set of chemical compounds (nodes) and pathways (edges), and the merging process may correspond to reminiscent tinkering processes such as bridging pathways between pre-existing nodes, enabling doubling pathways. These processes may positively or negatively contribute to increase robustness against mutations or give divergence and become a new pathway. The assumption of modular elements in our model is supported because module structures are found in many biological networks [23,45].

The original PA mechanism is exactly proportional to the number of edges of the node [16]. The FD-PA mechanism can consider additional or extra fitness of nodes. Rozenfeld *et al.* [34] and Zhu *et al.* [35] introduce into nodes the additional factors aging and a radius from a node in two-dimensional space. Because the additional factors are determined on the basis of a node's property, it is easy for us to infer that the effects from the factor is considered to be underlie physicochemical properties of biomolecules such as mass, length, affinity, structural stabilities, and contents of bases or amino acids in a biomolecule. Together with the idea increasing fitness during the growth process reflects a certain variation of node's properties with time, and is relevant to specification and/or fixation process after duplication and divergence.

11 Summary

We have presented analytical and numerical results verifying that our model produces the three remarkable statistical properties that are widely shared in biological network structures.

Scale-free connectivity — The degree distribution of our model follows power law, indicating the scale-free property; thus $P(k) \propto k^{-\gamma}$. The degree exponent γ can be predicted by Eq. (11), and demonstrates the range $2 < \gamma < \infty$ which includes the range of the degree exponents that are observed in biological

networks $2 < \gamma < 3$ [10,12,18]. The degree exponents take smaller values, meaning that the degree distribution becomes flat for the larger m/a and ξ .

Hierarchical modularity — The clustering spectrum of our model follows the power law with the exponent -1 ; hence $C(k) \propto k^{-1}$, reflecting the hierarchical modularity [24,25]. The power-law spectrum is nearly independent of ξ and m/a . The modularity of the whole network is higher for large ξ and smaller m/a . The hierarchy manifests growing mechanism by merging modules.

Disassortativity — The degree-degree correlation of our model follows the power law; therefore $\bar{k}_{nn}(k) \propto k^\nu$, reflecting the disassortativity. The exponent ν has the range $-1 < \nu < 0$ which includes the range of the exponents that are seen in biological networks [13,14,29]. The result indicates that the FD mechanism with positive ξ is critical to reproduce this property. In addition, the disassortativity is sensitive to ξ and m/a .

Acknowledgement

This work was partially supported by Grant-in-Aid No.18740237 from MEXT (JAPAN).

References

- [1] T.I. Lee *et al.*, Transcription regulatory networks in *Saccharomyces cerevisiae*, Science 298 (2002) 799.
- [2] I. Farkas *et al.*, The topology of transcription regulatory network in the yeast, *Saccharomyces cerevisiae*, Physica A 318 (2003) 601.
- [3] N. M. Luscombe *et al.*, Genomic analysis of regulatory network dynamics reveals large topological changes, Nature 431 (2004) 308.
- [4] P. Uetz, *et al.*, A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*, Nature 403 (2000) 623.
- [5] T. Ito, *et al.*, A comprehensive two-hybrid analysis to explore the yeast protein interactome, Proc. Natl. Acad. Sci. USA 98 (2001) 4569.
- [6] L. Giot *et al.*, A protein interaction map of *Drosophila melanogaster*, Science 302 (2003) 1727.
- [7] J.-F. Rual *et al.*, Towards a proteome-scale map of the human protein-protein interaction network, Nature 437 (2005) 1173.

- [8] H. Jeong *et al.*, The large-scale organization of metabolic networks, *Nature* 407 (2000) 651.
- [9] H. Ma, A.-P. Zeng, Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms, *Bioinformatics* 19 (2002) 270.
- [10] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* 74 (2002), 47.
- [11] A.-L. Barabási, Z. N. Oltvai, Network biology: understanding the cell's functional organization, *Nat. Rev. Genet.* 5 (2004) 101.
- [12] R. Albert, Scale-free networks in cell biology, *J. Cell Sci.* 118 (2005) 4947.
- [13] V. Colizza, A. Flammini, A. Maritan, A. Vespignani, Characterization and modeling of protein-protein interaction networks, *Physica A* 352 (2005) 1.
- [14] K.-I. Goh, B. Kahng, D. Kim, Graph theoretic analysis of protein interaction networks of eukaryotes, *Physica A* 357 (2005) 501.
- [15] A.-L. Barabási, R. Albert, Emergence of Scaling in Random Networks, *Science* 286 (1999) 509.
- [16] A.-L. Barabási, R. Albert, H. Jeong, Mean-field theory for scale-free random networks, *Physica A* 272 (1999) 173.
- [17] P. Erdős, A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* 5 (1960), 17.
- [18] J. F. F. Mendes and S. N. Dorogovtsev, *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford Press, New York, 2003.
- [19] S. Bornholdt, H. G. Schuster, *Handbook of Graphs and Networks: From the Genome to the Internet*, WILEY-VCH, Berlin, 2003.
- [20] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach*, Cambridge University Press, Cambridge, U.K., 2004.
- [21] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-worlds' networks, *Nature* 393 (1998) 440.
- [22] L. H. Hartwell, J. J. Hopfield, S. Leibler, A. W. Murray, From molecular to modular cell biology, *Nature* 402 (1999) C47.
- [23] C. v. Mering *et al.*, Genome evolution reveals biochemical networks and functional modules, *Proc. Natl. Acad. Sci. USA* 100 (2003) 15428.
- [24] E. Ravasz *et al.*, Hierarchical Organization of Modularity in Metabolic Networks, *Science* 297 (2002) 1551.
- [25] E. Ravasz, A.-L. Barabási, Hierarchical organization in complex networks, *Phys. Rev. E* 67 (2003) 026112.

- [26] A. Barrat, R. Pastor-Satorras, Rate equation approach for correlations in growing network models, *Phys. Rev. E* 71 (2005) 36127.
- [27] K. Takemoto, C. Oosawa, Evolving networks by merging cliques, *Phys. Rev. E* 71 (2005) 046116.
- [28] R. Pastor-Satorras, A. Vázquez, A. Vespignani, Dynamical and correlation properties of the internet, *Phys. Rev. Lett.* 87 (2001) 258701.
- [29] S.-H. Yook, F. Radicchi, H. Meyer-Ortmanns, Self-similar scale-free networks and disassortativity, *Phys. Rev. E* 72 (2005) 045105(R).
- [30] M. E. J. Newman, Assortative mixing in networks, *Phys. Rev. Lett.* 89 (2002) 208701.
- [31] A. Barrat, M. Barthélemy, A. Vespignani, Modeling the evolution of weighted networks, *Phys. Rev. E* 70 (2004) 066149.
- [32] G. Bianconi, A.-L. Barabási, Competition and multiscaling in evolving networks, *Europhys. Lett.* 54 (2001) 436.
- [33] A. Barrat, M. Barthélemy, A. Vespignani, Weighted evolving networks: Coupling topology and weight dynamics, *Phys. Rev. Lett.* 92 (2004) 228701.
- [34] A. F. Rozenfeld, R. Cohen, D. ben-Avraham, S. Havlin, Scale-free networks on lattices, *Phys. Rev. Lett.* 89 (2002) 218701.
- [35] H. Zhu, X. Wang, J.-Y. Zhu, Effect of aging on network structure, *Phys. Rev. E* 68 (2003) 056121.
- [36] R.V. Solé, S. Valverde, Information theory of complex networks: On evolution and architectural constraints, *Lect. Notes Phys.* 650 (2004) 189.
- [37] H. Salgado *et al.*, RegulonDB (version 4.0): transcriptional regulation, operon organization and growth conditions in *Escherichia coli* K-12, *Nucleic Acids Res.* 32 (2004) D303.
- [38] H. Jeong, S. Mason, A.-L. Barabasi, Z. N. Oltvai, Centrality and lethality of protein networks, *Nature* 411 (2001) 11.
- [39] S. Li *et al.*, A map of the interactome network of the metazoan *C. elegans*, *Science* 303 (2004) 540.
- [40] A. Rzhetsky, S. M. Gomez, Birth of scale-free molecular networks and the number of distinct DNA and protein domains per genome, *Bioinformatics* 17 (2001) 988.
- [41] R. Alves, R. A. G. Chaleil, M. J. E. Sternberg, Evolution of Enzymes in Metabolism: A Network Perspective, *J. Mol. Biol.* 320 (2002) 751.
- [42] E. Eisenberg, E. Y. Levanon, Preferential attachment in the Protein Network Evolution, *Phys. Rev. Lett.* 91 (2003) 138701.
- [43] R. Pastor-Satorras, E. Smith, R. V. Solé, Evolving protein interaction networks through gene duplication, *J. Theor. Biol.* 222 (2003) 199.

- [44] S. A. Teichmann, M. M. Babu, Gene regulatory network growth by duplication, Nat. Gene. 36 (2004) 492.
- [45] S. Wuchty, Z. N. Oltvai, A.-L. Barabási, Evolutionary conservation of motif constituents in the yeast protein interaction network, Nat. Gene. 35 (2003) 176.